

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-368787

(43)Date of publication of application : 20.12.2002

(51)Int.Cl.

H04L 12/56

(21)Application number : 2001-177011

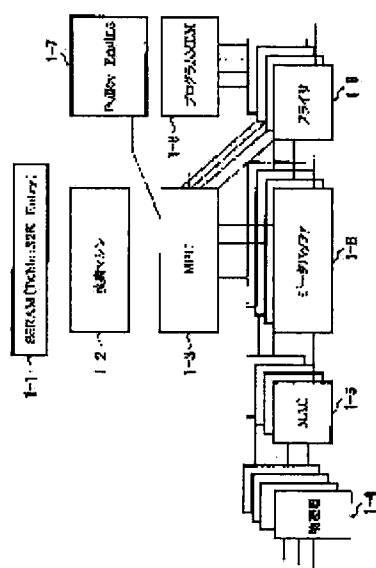
(71)Applicant : FUJITSU LTD

(22)Date of filing : 12.06.2001

(72)Inventor : ISHIDA NORIHIRO

(54) EXPLICIT PATH DESIGNATION RELAY DEVICE

本発明は、ネットワーク上の特定の経路指定を制御する装置に関する。



(57)Abstract:

PROBLEM TO BE SOLVED: To provide an explicit path designation relay device for supporting an explicit routing of an IPv6 packet that quickly switches a congested link to a bypass so as to relay traffic in a distributed way.

SOLUTION: The relay device monitors a traffic state of its own line interfaces 1-4 and informs other surrounding relaying devices of statistic information. A processor 1-3 discriminates whether or not bypass relay is made on the basis of the traffic statistic information notified from the other relaying devices and path information stored in a memory 1-1 of the own relaying device and sets an explicit path to designate a bypass relay path. In this case, explicit path designation information retrieved by a retrieval machine 1-2 is additionally inserted to a path control header at an entrance edge of a domain, and an inserted position pointer of the inserted path and a

relay designation stage number are described in a field in an extension header. Then the attached explicit path designation information is deleted at an exit edge of the domain.

* NOTICES *

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.*** shows the word which can not be translated.

3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1]Repeating installation which is the repeating installation which constitutes a network which relays data via two or more repeating installation, and possesses an explicit routing control means for specifying a course of alternate routing, comprising:

A means to supervise traffic conditions of a line interface which self-repeating installation accommodates, and to collect as statistical information.

A means to notify this statistical information to other surrounding repeating installation.

A means to judge whether alternate routing is carried out based on network path information held in this statistical information and self-repeating installation which were notified from other repeating installation.

Have an explicit routing control means for specifying a course of alternate routing, and said explicit routing control means, When the repeating installation concerned is located in entrance edge of a network domain, Addition insertion of the repeating-installation address list as explicit routing information used for a path control header for explicit routing only in this domain is carried out, A means to describe an added insertion point pointer and a relay specification number of stages of a course in the field in an extended header of Internet Protocol, A means to delete repeating-installation address list information as explicit routing information by which addition insertion was carried out in entrance edge repeating installation when the repeating installation concerned is located in exit edge of a network domain.

[Claim 2]The explicit routing repeating installation comprising according to claim 1:

A means to recognize repeating installation under suspension in said network domain.

A means which chooses an address of repeating installation under [out of a repeating-installation address of said path control header] employment, and is replaced in a destination address of an IP header into a repeating-installation address list described as said explicit routing information when an address of repeating installation under this suspension exists.

[Claim 3]The explicit routing repeating installation comprising according to claim 1 or 2:

A means by which message transfer capacity on a relay path of inputted packet data detects a priori whether it is size from message transfer capacity on an input interface.

A means to control not to perform addition insertion of a repeating-installation address list to said path control header when message transfer capacity of an output interface is smaller than message transfer capacity of an input interface.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention]In the system etc. by which especially this invention performs routing and relay of an IPv6 (Internet Protocol version 6) packet about explicit routing repeating installation, So that the IP packet of the user who should act as intermediary may be relayed via the course determined by the path control mechanism by autonomous and delivered control, and a different course (explicit course), The path control header of IPv6 is set up, a user IP packet is distributed and relayed in a network, and the scope and expansion of a traffic engineering system which cancel the maldistribution of the stagnation traffic in a network are planned.

[0002]In recent years, intensification of increase of the data traffic in use of the Internet is enhanced, and the improved efficiency of the network repeating installation which constitutes the Internet is called for. Expansion-izing of relay capacity and the realization of quality-of-service (QoS) guarantee formal ** which carries out priority relay about a specific IP packet are demanded especially.

[0003]with improvement-ization of the throughput of the user datum in the whole network including the packet relay of the best effort type by the "traffic engineering art" which supports the art corresponding to these demands. Realization of the optimum usage art of a repeating-installation resource covering the whole network is expected.

[0004]

[Description of the Prior Art]Drawing 11 shows the functional constitution of the repeating installation which performs the conventional explicit routing. This repeating installation, In a multiprotocol-label-switching (MPLS) system. Explicit routing protocols provided, such as CR-LDP (Constraint base -Label distribution Protocol) or RSVP (Resource Reservation Protocol). It uses and load sharing relay is performed.

[0005]In drawing 11, a circuit interface part and 11-2 11-1 A shortest route calculation control section, A channel information database (DB) and 11-4 11-3 A label distribution control section (treating part which gives the number-of-times display object of an explicit path change to a message), 11-5 shows an explicit route determination means, and 11-6 shows an statistical information receiving processing part (the function to judge the congestion of network repeating installation is included.), respectively.

[0006]The route determination mechanism by which the protocol of IPv6 was automated originally. For courses other than the course computed by (for example, OSPF (Open Shortest Path First) etc.). The path control header function (this header field is called Routing header.) for carrying out alternate routing of the IPv6 packet which is an user datum can be added.

[0007]Originally, to the commo data which a transmitting agency transmits, this path control header is added as an extended header in order to pass courses other than a default course. However, it is not performed conventionally that repeating installation changes the course for transmitting an user datum arbitrarily autonomously, i.e., carry out adding processing of the path control header (Routing header) by judgment of repeating installation.

[0008]Drawing 12 shows the example of the load sharing relay in a multiprotocol-label-switching (MPLS) network of operation. The label-switching router (LSR) of the entrance (Ingress) edge device with which 12-1 has a circuit interface part in drawing 12, Similarly 12-2 The label-switching router (LSR) of an exit

(Egress) edge device, They are the default route as which 12-3 was determined by the shortest route calculation control section, and the engineering route (alternate route for load sharing) which calculated 12-4,12-5 from the channel information database (DB), and was set up.

[0009]Multiprotocol label switching (MPLS) is mainly used for explicit routing from the former, As opposed to the data packet (either IPv4 or IPv6) from a user, The label equivalent to the connection identifier of a data link layer or it is given, Relay processing based on a label is performed by setting up the transmission line where an entrance (Ingress) edge device to the exit (Egress) edge device continued on the link between each label-switching router (LSR).

[0010]

[Problem(s) to be Solved by the Invention]The explicit routing control by the label switching path (LSP) tunnel in course reconstruction of the conventional multiprotocol label switching (MPLS), Between each label-switching router (LSR), path setting protocols, such as CR-LDP and RSVP, will be exchanged mutually, and the repeating installation of each will perform explicit routing control autonomously.

[0011]When the method which performs the path change in specific repeating installation here by the table management in the repeating installation which the protocol message in which the repeating installation and an explicit routing object (ERO) are contained passes is adopted, After making an autonomous judgment, the repeating installation which received congestion information needs to transmit and receive the message of this path setting protocol, in order to set up another alternate route.

[0012]And in transmission and reception of the message of the above-mentioned path setting protocol, in order to wait for the return of a completion response to the protocol which transmitted, the delay beyond fixed time certainly occurs by the completion of setting out of a new detour path. Therefore, it is difficult to set up a new alternate route timely and to change a data relay course in the explicit routing control by transmission and reception of a protocol message.

[0013]This invention, without changing the route table form in repeating installation, etc. entirely, Without performing path setting processing by transmission and reception of the protocol message of path setting control between repeating installation, Therefore, the prompt course change without delay (time lag) of an overhead and processing is realized, and the path which bypasses the specific link which carried out congestion is reconstructed, and the whole network is covered and it aims at providing the explicit routing repeating installation which can distribute and relay traffic.

[0014]When the repeating installation of the address described by the path control header has stopped operation according to the obstacle etc. in the explicit relay control by a path control header, An IPv6 packet will be discarded, without returning transmitting [the error message for which destination address attainment is improper] origin, or returning no error message, and an user datum will be again transmitted by the resending control of a host device into a network.

[0015]Now, the fault of relay processing needs to cope with [making an end host device perform management to the fault of the relay processing in a network, and] it with the repeating installation in a network essentially. An object of this invention is to provide the explicit routing repeating installation which can continue relay processing as it is without returning an error message etc., when processing of a path control header which makes a description change is performed actively and the repeating installation under suspension exists.

[0016] If processing which the repeating installation in the middle of a course tends to insert path control information in a path control header field autonomously, and is going to change the relay path of an IPv6 packet is performed, the overall length of a user IPv6 packet is made increased. A transmitting agency especially at the beginning and a host device, When the IPv6 packet of packet length according to the message transfer capacity (MTU) for which it searched with the protocol (PMTU) which searches for the message transfer capacity (MTU:Maximum Transfer Unit) under path is transmitted as an user datum, In the repeating installation in the middle of a course, when the message transfer capacity (MTU) of an output line interface is smaller than the message transfer capacity (MTU) of a line input interface, it is necessary to prevent the relay processing accompanied by insertion of a path control header.

[0017] When the message transfer capacity (MTU) of an output line interface is smaller than the message transfer capacity (MTU) of a line input interface, this invention, The explicit routing repeating installation which does not perform relay processing accompanied by insertion of the path control header to which the overall length of a packet is made to increase is provided.

[0018]

[Means for Solving the Problem] A path control function to specify a transit route by header information of an IPv6 packet is used for this invention, Repeating installation specifies a different course from an original course autonomously, and, on the other hand, with repeating installation of which this path change processing is canceled. Deletion of only a repeating-installation address of an inserted transit route for traffic distributions is performed, and timely traffic distribution processing in a traffic engineering domain is performed.

[0019] Namely, explicit routing repeating installation of this invention is repeating installation which constitutes a network which relays data via repeating installation of (1) plurality, In repeating installation possessing an explicit routing control means for specifying a course of alternate routing, A means to supervise traffic conditions of a line interface which self-repeating installation accommodates, and to collect as statistical information, A means to judge whether alternate routing is carried out based on network path information held in this statistical information and self-repeating installation it was notified from other repeating installation that were means to notify this statistical information to other surrounding repeating installation, Have an explicit routing control means for specifying a course of alternate routing, and said explicit routing control means, When the repeating installation concerned is located in entrance edge of a network domain, Addition insertion of the repeating-installation address list as explicit routing information used for a path control header for explicit routing only in this domain is carried out, When a means to describe an added insertion point pointer and a relay specification number of stages of a course in the field in an extended header of Internet Protocol, and the repeating installation concerned are located in exit edge of a network domain, A means to delete repeating-installation address list information as explicit routing information by which addition insertion was carried out in entrance edge repeating installation is provided.

[0020] (2) -- a means to recognize repeating installation under suspension in said network domain, [have and] When an address of repeating installation under this suspension exists in a repeating-installation address list described as said explicit routing information, an address of repeating installation under [out of a repeating-installation address of said path control header] employment is chosen, It has a means changed

to a destination address of an IP header.

[0021] Message transfer capacity on a relay path of packet data which carried out (3) inputs, A means to detect a priori whether it is size from message transfer capacity on an input interface is provided, When message transfer capacity of an output interface is smaller than message transfer capacity of an input interface, it has a means to control not to perform addition insertion of a repeating-installation address list to said path control header.

[0022]

[Embodiment of the Invention] The functional block of the explicit routing repeating installation by this invention is shown in drawing 1. The explicit routing repeating installation by this invention is provided with a means to add the address list of the repeating installation on a specification course to the path control header field for explicit routing, and is provided with the function to process the insert/delete of this path control header and to transmit an IPv6 packet.

[0023] Judgment and processing of the insert/delete of a path control header are performed by the microprocessor (MPU) 1-3. The insert/delete of the address list of repeating installation is performed according to the local regulation applied only to each repeating installation in a domain. The repeating-installation address list which specifies an explicit relay path is stored in the routing table (SSRAM) 1-1.

[0024] After the termination of a user's IPv6 packet inputted from the circuit of four ports is carried out by the MAC (Media Access Control) treating part 1-5 through the physical layer interface part 1-4, it is stored in the data buffer 1-6. The microprocessor (MPU) 1-3 about a user's IPv6 packet stored in the data buffer 1-6. Header inspection processing is performed according to the processing program stored in the program memory 1-8, required header information is read, and processing of the insert/delete of a path control header is judged.

[0025] And the microprocessor (MPU) 1-3 extracts a destination address from header information, and in order to determine an appearance side way, it transmits this destination address to the search machine 1-2. And when the response which contains the memory address pointer with which the additional address list to the path control header was stored as the search results is returned, the microprocessor (MPU) 1-3 performs processing which inserts the contents of this address list in a user's IPv6 packet.

[0026] The program memory 1-8 is used as data and the memory for commands, a request, the memory for status, a memory for statistical information collection, etc. The polish engine 1-7 provides an intelligent integrated packet classification function, and the slicer 1-9 has a conversion process function of an IP packet and an ATM cell.

[0027] Drawing 2 shows the process flow which inserts above-mentioned path control header information in a user's IPv6 packet. Drawing 3 shows the process flow which deletes the inserted path control header information in the repeating installation by the side of appearance. The pointer which shows drawing 4 an IPv6 header format and shows drawing 5 the insertion point of path control header information is shown.

[0028] When the processing which inserts path control header information in an IPv6 packet is explained with reference to drawing 2, the router of the entrance edge repeating installation of an IPv6 domain, While receiving traffic statistical information, routing table is searched with the destination address (DA) of a user's IPv6 packet inputted from the circuit (2-1).

[0029]And it is judged whether the data in which those with LRTE (Loose Route Traffic Engineering: traffic engineering which thins out and sets up repeating-installation address of specification course) execution are shown is set as the above-mentioned search results (2-2). The output interface number, the repeating-installation address through which it should pass, etc. are set up as this data.

[0030]When it judges whether the packet which received has a path control header when the above-mentioned data is set up (2-3) and there is no path control header, the path control header inserted newly is loaded to a memory (2-4). And the address of the repeating installation which should be passed is loaded from search results, and it inserts in a path control header (2-5).

[0031]Next, the pointer in which the insertion point of the inserted repeating-installation address is shown is edited (2-6). A path control header and an insertion point pointer are inserted in a receive packet (2-7), the following header field value of an IPv6 header is copied to the following header field of a path control header, and the value which shows a path control header to the following header field of an IPv6 header is written in after that (2-8). And a receiving IPv6 packet is transmitted based on the search results of routing table (2-9). (forwarding)

[0032]Next, when the deletion of path control header information is explained with reference to drawing 3, the router of the exit edge repeating installation of IPv6 / MPLS domain, If the IPv6 packet from the domain under traffic engineering (TE) execution by a path control header is received (3-1), It judges whether the repeating-installation address insertion point pointer to a path control header is 0 (3-2), and when it is except zero, the address list deletion from a repeating-installation address insertion point pointer or deletion of a path control header is judged (3-3).

[0033]When it is address list deletion, the address list inserted in the path control header is deleted by the insertion point pointer information to a path control header (3-4). When it is deletion of a path control header, by the insertion point pointer information to a path control header. The following header field value given to the path control header which deleted the path control header itself inserted in the path control header (3-5), and was deleted to the following header field in the header of the preceding paragraph of a path control header is written in (3-6). And the part which deleted path control header information, and payload length are subtracted, it writes in an IPv6 header (3-7), and the data transfer (forwarding) to the following course is started (3-8).

[0034]The basic header in which the header format of IPv6 includes a version, a traffic class (priority), a flow label, payload length, the following header, hop restrictions, a transmission source address, and a destination address as shown in drawing 4, The extended header containing a HOPPUBAI hop option header, an address option header, a routing header, a fragmentation header, an authentication header, an upper layer header, etc. is comprised.

[0035]The relay station address of a transit route is written in the routing header of the above-mentioned extended header. The format of this routing header is shown in drawing 5. The figure shows the format of the routing header of Type 0, and the routing header is equipped with the field of the following header, extended header length, a routing type, a ** segment, the reserve (Reserved) field, and each repeating-installation address of a transit route.

[0036]The field of the following header is an 8-bit selector, the type of the header which continues immediately after a routing header is shown, and the same value as the protocol field in IPv4 packet is used.

The field of extended header length expresses the length of 8 octet units of a routing header for the numerals-less integer of 8 bits (however, the amount of eight top octets do not contain). The value of the extended header length of the routing header of Type 0 becomes equal to the address numbers in a routing header twice the value of repeating installation.

[0037]A routing type value (in this case, "0") is written in the routing type field at 8 bits. The number of the intermediate nodes (repeating installation) clearly shown as going for the numerals-less integer of 8 bits by the time it reaches the remaining numbers of segments on a course, i.e., a final destination, is written in the field of a ** segment.

[0038]In a 32-bit information field, the reserve (Reserved) field is initialized by "0" at the time of transmission, and it is ignored in a receiver. Each repeating-installation address [1] of a transit route -- [n] is an address of the 128 bit length of each repeating installation with which a number was assigned from 1 to n.

[0039]A 1st embodiment of this invention inserts the format of the following control fields in the reserve (Reserved) field as an example as a pointer in which the insertion point of the repeating-installation address to a path control header is shown.

- (1) Define and insert the field of the SIMM (shim) header format in a multiprotocol-label-switching (MPLS) method in this reserve (Reserved) field.
- (2) Make the number of stages of a SIMM (shim) header into one step.
- (3) If the data of the reserve (Reserved) field is not "0", it will be judged that the SIMM (shim) header is inserted as significant data.
- (4) In this case, divide the label field (20 bits) of a SIMM (shim) header into two octets and a 4-bit subfield, and describe the range specification of the insertion route list at the time of the traffic engineering by a path control header.

[0040]A SIMM (shim) header format is shown in drawing 6. This SIMM (shim) header format is specified by RFC3032Jan.2001. A SIMM (shim) header format comprises 32 bits of the 20-bit label (Label) field, the experiment use (Exp) field of a triplet, the 1 bit bottoms tuck (S) field, and the 8-bit life time (TTL) field.

[0041]In the 1st 8-bit subfield (sub-label-1) in the label (Label) field, the pointer in which the position which inserts the contents of a path control header for traffic engineering is shown is described. In the 2nd 8-bit subfield (sub-label-2), the pointer in which the number which inserts the contents of a path control header for traffic engineering is shown is described. The 4-bit subfield (sub-label-3) in the label (Label) field is a reserve (Reserved).

[0042]Thus, a user's IPv6 packet is detoured to another course by including the control mechanism of a path change in the path control header of an IPv6 packet, without a path setting control protocol performing path setting processing. In this case, in order to notify explicit routing information to other repeating-installation groups, a 1st embodiment of this invention, The insertion point pointer and relay specification number of stages of a course (repeating-installation address list) which were newly added all over the reserve (Reserved) field of the path control header which is the field of the TLV (Type/Length/Value) group contained in an extended header field are described.

[0043]In this case, in the repeating installation which is not supporting the control mechanism of this path change, If there are some which perform processing which clears the reserve (Reserved) field of a path

control header to "0", it will become impossible for the repeating installation of exit edge to perform deletion of the newly added course, and fault will occur in operation.

[0044]Namely, although it indicated that addition insertion of the repeating-installation address list is carried out by the reserve (Reserved) field in a path control header and a path change is controlled, Originally the reserve (Reserved) field, In [are the field secured for future use and] mounting of other IPv6 packet relay mechanisms, Being used for various judgment, processings, etc. which carry out the zero clear of this reserve (Reserved) field, and relay it, or alarm processing will be carried out as malfunction detection if those contents are checked and it becomes except zero is expected.

[0045]Then, the repeating installation on a course can be specified as a header which should always be inspected, field regulation can be newly added to the hop by hop option field which is a general-purpose header field usable to other uses, and insertion of a repeating-installation address list can be displayed. It is a part of field value which was specified as a hop by hop option field, and it has not been specified other than it.

[0046]As a value of this specified portion, a triplet eye from an option type higher rank, As this option data is to the final destination of a packet, it is a bit which specifies whether it is correctable, and the fault by the above-mentioned zero clear can be solved by setting this bit as the value (0) which cannot change option data on the way. However, the load of the insertion deletion of a HOPPUBAI hop option header increases.

[0047]However, in order to avoid the fault by reserve (Reserved) field use, Instead of the reserve (Reserved) field of a path control header, as a 2nd embodiment of this invention, The pad N option function of the HOPPUBAI hop option header of IPv6 protocol standard regulation is extended, and it carries out to making the newly added insertion point display pointer of a course, and the number specification of relay stages display.

[0048]The field definition of the pad N option (alignment requirement : none) in IPv6 protocol standard specification is explained below. A pad N option is used for inserting padding of two or more octets in the option area of a header. The format of the pad N option (alignment requirement : none) field is shown in (a) of drawing 7. In padding of N octet, the value of (N-2) is described in the option data length field, and "0" values of a octet are described by the option data field (N-2).

[0049]According to regulation of IPv6 protocol standard specification, must process a series of options in a header strictly in the order which appeared in the header, for example, the specific option in a header is looked for, before processing all the options preceded with it — the option — it must not process — carrying out — having — **** . Pad 1 option is used for inserting padding of one octet in the option area of a header, and when padding exceeding one octet is required, the pad N option shown here is used [rather than] using two or more pad 1 options.

[0050]An option type identifier has the above role that only identifies the specific type of an option, and it is coded by the top 2 bits so that it may specify how it treats, when the IPv6 node under processing cannot recognize the option type. The meaning of the numerals of 2 bits of option type higher ranks is specified as shown in (b) of drawing 7, and the pad 1 above-mentioned option and a pad N option follow regulation of the numerals "00" of the table of drawing 7 (b). If repeating installation receives an IPv6 packet with a HOPPUBAI hop option header, the option type is inspected, and if it is the pad 1 above-mentioned option or a pad N option, after performing processing corresponding to it, the next extended header processing will be

performed.

[0051]The SIMM (shim) header of a multiprotocol-label-switching (MPLS) method is inserted in the option data part of the above-mentioned pad N option as an example here about the insertion point pointer of an additional course (repeating-installation address list) and the description of a relay specification number of stages in this invention. By referring to the option data length field, the number of stages of a SIMM (shim) header can distinguish a SIMM (shim) number of stages, if a number of times of four octets are calculated. [0052]"00000XXX" is used for an option type value, and if option data is not "0", it will be judged that the SIMM (shim) header is inserted as significant data. In this case, the label field (20 bits) of a SIMM (shim) header is divided into two octets and a 4-bit subfield, and the range specification of the insertion route list at the time of the traffic engineering by a path control header is described.

[0053]When an IPv6 packet with a HOPPUBAI hop option header is received, the insertion point pointer of path control header information is described in the form added to this. When a user IPv6 packet without a HOPPUBAI hop option header is received, a HOPPUBAI hop option header is set up newly.

[0054]In this case, it is necessary to rewrite to the display of the purport that it is a HOPPUBAI hop option header in the following header field in front of the path control header to insert. In this pad N option, the insertion point information pointer of a path control header as shown in drawing 6 is described. The format of this position information pointer can divert the SIMM (shim) header format shown in drawing 6 as an example.

[0055]Next, the example of a process flow when the repeating installation (router) of the following hop is carrying out suspension into the domain of an IPv6 relay system is shown in drawing 8. Since each repeating installation is usually operating the dynamic routing protocol represented by OSPF (Open Shortest Path First), it can detect whether the adjoining repeating installation is [*****] under employment.

[0056]Then, if the IPv6 packet from the traffic engineering implementation domain by a path control header is received (8-1), It is judged whether the repeating installation of the address described by the header of the IPv6 packet is operated normally (8-2), Though the channel information to the repeating installation was described by the path control header (routing table) when operation was suspended according to the obstacle etc., Processing which exchanges the next relay address and the destination address of an IPv6 basic header which are described by the path control header is performed (8-3), IPv6 packet transmission to the following course is performed (8-4), and the relay of a user's IPv6 packet is continued.

[0057]And in a stage while being the relay processing which a host device cannot know thus, the path control header from the first which repeating installation set up the value of the path control header autonomously, and was actively given to the user IPv6 packet for the congestion relaxation in a network -- or insertion deletion of a path control header is completely performed newly. In this case, the repeating installation of the direct previous hop of what caused the obstacle and has suspended relay operation with the repeating installation in that network domain, If the repeating installation under stop is recognized (distinction by transmission and reception of a Keep alive packet is possible), in the repeating installation, address list exchange processing of a path control header will be carried out, and a user IPv6 packet will be relayed towards the next destination address (course under employment).

[0058]Such [usually] processing is not performed about the path control header which the user set up. The error message for which address attainment is improper will be returned to a transmission source device,

and will be resent by processing of a transmission source device. However, since a path control header is created and inserted with the repeating installation in the middle of a course in this invention, The relay of a user IPv6 packet can be continued by performing autonomously processing replaced by a destination address while employing the contents of the path control header, without leaving resending control to a transmission source host device, without hanging a processing load on a transmission source host device.

[0059]Next, the processing which performs insertion control of path control information based on the message transfer capacity of an input interface and an output interface is explained with reference to drawing 9. The explicit routing repeating installation by this invention must be conscious of the difference in the message transfer capacity (MTU) size of a relay path.

[0060]For this reason, addition insertion of path control information is controlled based on the message transfer capacity (MTU) size of a relay path by inspecting the contents of the "Interface" field which usually has a statement in the routing table which repeating installation holds. In the "Interface" field, specifically, for example Ethernet (registered trademark), Since interface classification (link classification), such as FDDI (Fiber Distributed Data Interface) and ATM (Asynchronous Transfer Mode), is described, With reference to this, the size relation of message transfer capacity (MTU) size can be distinguished.

[0061]The label-switching router (LSR) of the domain entrance edge in IPv6 / multiprotocol-label-switching (MPLS) system, The size of the message transfer capacity (MTU) of the ON side interface is compared with the size of the message transfer capacity (MTU) of the appearance side interface with reference to the interface classification (link classification) information on routing table, and size relation is judged.

[0062]As shown in drawing 9, the message transfer capacity (MTU) of the receiver interface of a user IPv6 packet is memorized first (9-1). Next, routing table is searched and the classification of the appearance side interface which performs load distribution processing in an IPv6 path control header is judged (9-2). And it is judged whether the size of the message transfer capacity (MTU) of the appearance side interface is larger than the message transfer capacity (MTU) of a receiver interface (9-3).

[0063]The flag which the size of the message transfer capacity (MTU) of the appearance side interface is smaller than a receiver interface, or forbids distribution by a path control header with the interface concerned when equal is set up (9-4). Repeating installation stops the processing which sets the additional address list to a path control header as the search results of routing table, when this inhibit flag is set up. As a result, on the interface concerned, routing control by a path control header is not performed.

[0064]The example of the holds information of routing table is shown in drawing 10. Here, the field described to be "Interface" shows interface classification (link classification), and it is distinguished from description of "ei31" in drawing 10 that interface classification (link classification) is Ethernet (registered trademark).

[0065]Then, the interface classification which received the user's IPv6 packet, The interface classification (the contents of the Interface field) produced by searching routing table is compared, and if the message transfer capacity (MTU) of an exit interface is not larger, it is made not to perform address adding processing to a path control header.

[0066](Additional remark 1) In the repeating installation possessing the explicit routing control means for being the repeating installation which constitutes the network which relays data via two or more repeating installation, and specifying the course of alternate routing, A means to supervise the traffic conditions of the line interface which self-repeating installation accommodates, and to collect as statistical information,

A means to judge whether alternate routing is carried out based on the network path information held in this statistical information and self-repeating installation it was notified from other repeating installation that were means to notify this statistical information to other surrounding repeating installation, Have an explicit routing control means for specifying the course of alternate routing, and said explicit routing control means, When the repeating installation concerned is located in the entrance edge of a network domain, Addition insertion of the repeating-installation address list as explicit routing information used for the path control header for explicit routing only in this domain is carried out, When a means to describe the added insertion point pointer and relay specification number of stages of a course in the field in the extended header of Internet Protocol, and the repeating installation concerned are located in the exit edge of a network domain, Explicit routing repeating installation possessing a means to delete the repeating-installation address list information as explicit routing information by which addition insertion was carried out in entrance edge repeating installation. (1)

(Additional remark 2) It has a means to recognize the repeating installation under suspension in said network domain, Explicit routing repeating installation given in the additional remark 1 provided with a means to replace the repeating-installation address of said path control header at the course to the repeating installation under employment into the repeating-installation address list described as said explicit routing information when the repeating installation under this suspension exists. (2)

(Additional remark 3) The message transfer capacity on the relay path of the inputted packet data, A means to detect a priori whether it is size from the message transfer capacity on an input interface is provided, The additional remark 1 having a means to control not to perform addition insertion of the repeating-installation address list to said path control header when the message transfer capacity of an output interface is smaller than the message transfer capacity of an input interface, or explicit routing repeating installation given in 2. (3)

(Additional remark 4) Explicit routing repeating installation given in the additional remark 1, wherein said explicit routing control means is provided with a means to describe the added insertion point pointer and relay specification number of stages of a course to the preliminary field in the path control header in the extended header of Internet Protocol.

(Additional remark 5) Said explicit routing control means, Explicit routing repeating installation given in the additional remark 1 provided with a means to describe the added insertion point pointer and relay specification number of stages of a course to the HOPPUBAI hop option header field in the extended header of Internet Protocol.

[0067]

[Effect of the Invention]By carrying out the insert/delete of the repeating-installation address list of the course to change to a path control header, when performing load sharing control which avoids concentration of traffic according to this invention, as explained above, Without changing the form of the routing table in repeating installation, etc. entirely, Without generating delay of the overhead by using protocols, such as LDP/CR-LDP for the path setting control between repeating installation or RSVP, and processing, The thing for carrying out distributed relay of the traffic load into a network to do for path change control becomes possible, without using a signaling function etc. for others.

[0068]In the domain under traffic engineering execution by this invention, Distributed relay of the traffic load

can be carried out without producing a problem entirely in other repeating installation by deleting an additional repeating-installation address with the exit edge of a domain, even if a repeating-installation address is temporarily added to a path control header.

[0069]In the device side which relays an IPv6 packet, it is possible to make an IPv6 user packet relay according to the course which usually performed the header check of the passage and was described by this path control header. As a result, the input load to the link which congestion had generated in the network can be decreased promptly, The communication throughput of the end user under Internet usage using this course can be increased seemingly, It can realize in the form which adds the prompt course change without a time lag to the IPv6 protocol processing indicated to the Internet draft or RFC, without generating inconsistency with standard protocol specification.

[0070]By changing the course to the repeating installation under this suspension to the course to the repeating installation under employment, when the repeating installation under suspension exists on the specification course of a path control header, Communication can be made to continue, without producing the interrupted activity of the relay processing by an error message [that it cannot reach to a destination host device] being returned, or an IPv6 packet being discarded, without returning no error message.

[Translation done.]

(19) 日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11) 特許出願公開番号

特開2002-368787

(P2002-368787A)

(43) 公開日 平成14年12月20日 (2002. 12. 20)

(51) Int.Cl.⁷

H 0 4 L 12/56

識別記号

1 0 0

F I

H 0 4 L 12/56

テーマコード* (参考)

1 0 0 B 5 K 0 3 0

審査請求 未請求 請求項の数 3 O L (全 14 頁)

(21) 出願番号 特願2001-177011(P2001-177011)

(22) 出願日 平成13年6月12日 (2001. 6. 12)

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72) 発明者 石田 憲弘

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(74) 代理人 100105337

弁理士 眞鍋 潔 (外3名)

Fターム(参考) 5K030 GA13 HA08 HB14 HD03 JA11

KA05 LB08 LE01 MA07 MB01

MB09

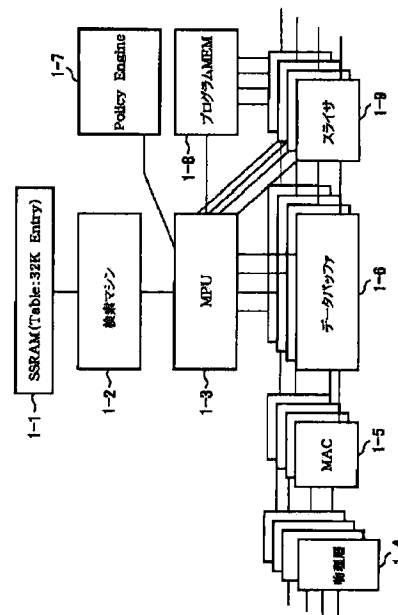
(54) 【発明の名称】 明示的経路指定中継装置

(57) 【要約】

【課題】 I P v 6 パケットの明示的ルーティングをサポートする明示的経路指定中継装置に関し、輻輳したリンクを迂回する経路へ速やかに切替え、トラフィックを分散中継する。

【解決手段】 自中継装置の収容する回線インタフェース1-4のトラフィック状況を監視し、該統計情報を周囲の他の中継装置に通知する。他の中継装置から通知されたトラフィック統計情報及び自中継装置のメモリ1-1に保有する経路情報を基に、プロセッサ1-3において迂回中継するか否かを判断し、迂回中継の経路を指定するための明示的経路を設定する。このとき、ドメインの入口エッジにおいて、検索マシン1-2により検索した明示的経路指定情報を経路制御ヘッダに付加挿入し、挿入した経路の挿入位置ポインタ及び中継指定段数を拡張ヘッダ内のフィールドに記述する。そしてドメインの出口エッジでその付加した明示的経路指定情報を削除する。

本発明による明示的経路指定中継装置の機能ブロック



【特許請求の範囲】

【請求項1】 複数の中継装置を経由してデータを中継するネットワークを構成する中継装置であって、迂回中継の経路を指定するための明示的経路設定制御手段を具備する中継装置において、

10 自中継装置の収容する回線インタフェースのトラフィック状況を監視し、統計情報として収集する手段と、周囲の他の中継装置に該統計情報を通知する手段と、他の中継装置から通知された該統計情報及び自中継装置に保有しているネットワーク経路情報を基に、迂回中継するか否かを判断する手段と、迂回中継の経路を指定するための明示的経路設定制御手段とを備え、

前記明示的経路設定制御手段は、当該中継装置がネットワークドメインの入口エッジに位置する場合、明示的経路指定用の経路制御ヘッダに、該ドメイン内でのみ使用する明示的経路指定情報としての中継装置アドレスリストを付加挿入し、付加した経路の挿入位置ポインタ及び中継指定段数を、インターネットプロトコルの拡張ヘッダ内のフィールドに記述する手段と、

20 当該中継装置がネットワークドメインの出口エッジに位置する場合、入口エッジ中継装置において付加挿入された明示的経路指定情報としての中継装置アドレスリスト情報を削除する手段と、を具備することを特徴とする明示的経路指定中継装置。

【請求項2】 前記ネットワークドメイン内の運用停止中の中継装置を認識する手段を備え、前記明示的経路指定情報として記述された中継装置アドレスリスト中に、該運用停止中の中継装置のアドレスが存在する場合、前記経路制御ヘッダの中継装置アドレスの中から運用中の中継装置のアドレスを選択して、IPヘッダの宛先アドレスに入れ替える手段を備えたことを特徴とする請求項1に記載の明示的経路指定中継装置。

【請求項3】 入力したパケットデータの中継経路上のメッセージ転送容量が、入力インタフェース上のメッセージ転送容量より大であるか否かを事前に検出する手段を具備し、出力インタフェースのメッセージ転送容量が入力インタフェースのメッセージ転送容量より小さい場合、前記経路制御ヘッダの中継装置アドレスリストの付加挿入を行わないように制御する手段を備えたことを特徴とする請求項1又は2に記載の明示的経路指定中継装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は明示的経路指定中継装置に関し、特に、IP v 6 (Internet Protocol version 6) パケットのルーティングと中継を行うシステム等において、中継すべきユーザのIPパケットを、自律分散制御による経路制御機構により決定される経路と異なる経路(明示的経路)を経由して中継するように、IP v 6の経路制御ヘッダを設定し、ユーザIPパケット

をネットワーク内において分散して中継し、ネットワーク内の滞留トラフィックの偏在を解消するトラフィックエンジニアリングシステムの適用範囲と機能拡張を図るものである。

【0002】近年、インターネットの利用におけるデータトラフィックの増大は激化の一途をたどり、インターネットを構成するネットワーク中継装置の性能向上が求められている。中でも中継容量の拡大化と、特定のIPパケットについて優先中継するサービス品質(QoS)保証型通信の実現化が要求されている。

【0003】これらの要求に合致する技術をサポートする“トラフィックエンジニアリング技術”によるベストエフォート型のパケット中継を含めたネットワーク全体におけるユーザデータのスループットの向上化と、ネットワーク全体に亘る中継装置リソースの最適利用技術の実現化が期待されている。

【0004】

【従来の技術】図11は従来の明示的ルーティングを行う中継装置の機能構成を示す。該中継装置は、マルチプロトコラベルスイッチング(MPLS)システムにおいて提供されているCR-LDP (Constraint base - Label distribution Protocol) 又はRSVP (Resource Reservation Protocol) 等の明示的ルーティングプロトコルを用いて負荷分散中継を行う。

【0005】図11において、11-1は回線インタフェース部、11-2は最短経路算出制御部、11-3は経路情報データベース(DB)、11-4はラベル分配制御部(明示的経路変更回数表示オブジェクトをメッセージに付与する処理部)、11-5は明示的経路決定手段、11-6は統計情報受信処理部(ネットワーク中継装置の輻輳を判断する機能を含む。)をそれぞれ示す。

【0006】また、IP v 6のプロトコルは、元来、自動化された経路決定メカニズム(例えば、OSPF (Open Shortest Path First) など)により算出された経路以外の経路に、ユーザデータであるIP v 6パケットを迂回中継させるための経路制御ヘッダ機能(このヘッダフィールドはRouting headerと呼ばれる。)を付加することができる。

【0007】この経路制御ヘッダは、本来、送信元が送信する通信データに対して、デフォルトの経路以外の経路を通過させるべく、拡張ヘッダとして付加されるものである。但し、ユーザデータを転送するための経路を中継装置が自律的に任意に変更すること、つまり、中継装置の判断により経路制御ヘッダ(Routing header)を追加処理することは従来行われていない。

【0008】図12はマルチプロトコラベルスイッチング(MPLS)ネットワークにおける負荷分散中継の動作例を示す。図12において、12-1は、回線インタフェース部を有する入口(Ingress)エッジ装置のラベルスイッチングルータ(LSR)、12-2は同じく

出口 (Egress) エッジ装置のラベルスイッチングルータ (LSR)、12-3は最短経路算出制御部により決定されたデフォルトルート、12-4、12-5は経路情報データベース (DB) から計算して設定したエンジニアリングルータ (負荷分散のための迂回経路) である。

【0009】従来からの明示的ルーティングは、マルチプロトコラベルスイッチング (MPLS) を主に用い、ユーザからのデータパケット (IPv4又はIPv6のいずれでも) に対し、データリンクレイヤのコネクション識別子又はそれに相当するラベルを付与し、各ラベルスイッチングルータ (LSR) 間のリンク上に、入口 (Ingress) エッジ装置から出口 (Egress) エッジ装置までの連続した伝送路を設定することにより、ラベルに基づく中継処理を行うものである。

【0010】

【発明が解決しようとする課題】従来のマルチプロトコラベルスイッチング (MPLS) の経路再構成におけるラベルスイッチングパス (LSP) トンネルによる明示的経路設定制御は、各ラベルスイッチングルータ (LSR) 間で、CR-LDPやRSVPといったパス設定プロトコルを相互に交換して、中継装置各々が自律的に明示的経路設定制御を実行することとなる。

【0011】ここで、特定の中継装置での経路変更を、その中継装置及び明示的経路指定オブジェクト (ERO) が含まれるプロトコルメッセージが通過する中継装置内のテーブル管理により行う方式が採用されている場合、輻輳情報を受信した中継装置は自律的な判断を行った後、別の迂回経路を設定するためにこのパス設定プロトコルのメッセージを送受信する必要がある。

【0012】そして上記パス設定プロトコルのメッセージの送受信において、送信したプロトコルに対する完了応答の返送を待たずするため、新規な迂回パスの設定完了までに一定時間以上の遅れが必ず発生する。そのため、プロトコルメッセージの送受による明示的経路設定制御では、タイムリーに新規迂回経路を設定してデータ中継経路の変更をすることが困難である。

【0013】本発明は、中継装置内の経路テーブル形式等を一切変更することなく、また、中継装置間でパス設定制御のプロトコルメッセージの送受によるパス設定処理を行うことなく、従ってオーバーヘッドや処理の遅延 (タイムラグ) のない速やかな経路切替えを実現し、特定の輻輳したリンクを迂回するパスを再構成し、ネットワーク全体に亘ってトラフィックを分散して中継することができる明示的経路指定中継装置を提供することを目的とする。

【0014】また、経路制御ヘッダによる明示的な中継制御において、経路制御ヘッダに記述されたアドレスの中継装置が障害等により運用を停止していた場合、宛先アドレス到達不可のエラーメッセージが送信元に返送され、或いは何のエラーメッセージも返送されずにIPv

6パケットが廃棄され、エンドホスト装置の再送制御により再度ユーザデータがネットワーク内に送信されることになる。

【0015】これでは、ネットワーク内の中継処理の不具合に対する対処をエンドホスト装置に実行させることとなり、中継処理の不具合は、本来ネットワーク内の中継装置で対処する必要がある。本発明は、能動的に経路制御ヘッダの記述変更する処理を行い、運用停止中の中継装置が存在する場合、エラーメッセージ等を返送することなくそのまま中継処理を継続することができる明示的経路指定中継装置を提供することを目的とする。

【0016】また、経路制御ヘッダフィールドに経路途中の中継装置が自律的に経路制御情報を挿入してIPv6パケットの中継経路を変更しようとする処理を行うと、ユーザIPv6パケットの全長を増加させてしまうことになる。特に、当初送信元のエンドホスト装置が、パス中のメッセージ転送容量 (MTU: Maximum Transfer Unit) を探索するプロトコル (PMTU) により探索したメッセージ転送容量 (MTU) に従ったパケット長のIPv6パケットをユーザデータとして送信している場合に、経路途中の中継装置において、出力回線インタフェースのメッセージ転送容量 (MTU) が入力回線インタフェースのメッセージ転送容量 (MTU) よりも小さいときには、経路制御ヘッダの挿入処理を伴う中継処理を防止する必要がある。

【0017】本発明は、出力回線インタフェースのメッセージ転送容量 (MTU) が入力回線インタフェースのメッセージ転送容量 (MTU) よりも小さいときには、パケットの全長を増加させてしまう経路制御ヘッダの挿入処理を伴う中継処理を行わない明示的経路指定中継装置を提供する。

【0018】

【課題を解決するための手段】本発明は、IPv6パケットのヘッダ情報により通過経路を指定する経路制御機能を用いて、本来の経路とは異なる経路を中継装置が自律的に指定し、一方、この経路変更処理を解除する中継装置では、挿入されたトラフィック分散用の通過経路の中継装置アドレスのみの削除処理を行って、トラフィックエンジニアリングドメイン内でのタイムリーなトラフィック分散処理を実行する。

【0019】即ち、本発明の明示的経路指定中継装置は、(1) 複数の中継装置を経由してデータを中継するネットワークを構成する中継装置であって、迂回中継の経路を指定するための明示的経路設定制御手段を具備する中継装置において、自中継装置の収容する回線インタフェースのトラフィック状況を監視し、統計情報として収集する手段と、周囲の他の中継装置に該統計情報を通知する手段と、他の中継装置から通知された該統計情報及び自中継装置に保有しているネットワーク経路情報を基に、迂回中継するか否かを判断する手段と、迂回中継

の経路を指定するための明示的経路設定制御手段とを備え、前記明示的経路設定制御手段は、当該中継装置がネットワークドメインの入口エッジに位置する場合、明示的経路指定用の経路制御ヘッダに、該ドメイン内でのみ使用する明示的経路指定情報としての中継装置アドレスリストを付加挿入し、付加した経路の挿入位置ポインタ及び中継指定段数を、インターネットプロトコルの拡張ヘッダ内のフィールドに記述する手段と、当該中継装置がネットワークドメインの出口エッジに位置する場合、入口エッジ中継装置において付加挿入された明示的経路指定情報としての中継装置アドレスリスト情報を削除する手段と、を具備するものである。

【0020】また、(2)前記ネットワークドメイン内の運用停止中の中継装置を認識する手段を備え、前記明示的経路指定情報として記述された中継装置アドレスリスト中に、該運用停止中の中継装置のアドレスが存在する場合、前記経路制御ヘッダの中継装置アドレスの中から運用中の中継装置のアドレスを選択して、IPヘッダの宛先アドレスに入れ替える手段を備えたものである。

【0021】また、(3)入力したパケットデータの中継経路上のメッセージ転送容量が、入力インタフェース上のメッセージ転送容量より大であるか否かを事前に検出する手段を具備し、出力インタフェースのメッセージ転送容量が入力インタフェースのメッセージ転送容量より小さい場合、前記経路制御ヘッダの中継装置アドレスリストの付加挿入を行わないように制御する手段を備えたものである。

【0022】

【発明の実施の形態】図1に本発明による明示的経路指定中継装置の機能ブロックを示す。本発明による明示的経路指定中継装置は、明示的経路指定用の経路制御ヘッダフィールドに指定経路上の中継装置のアドレスリストを付加する手段を備え、該経路制御ヘッダの挿入／削除の処理を行ってIPv6パケットを転送する機能を備えたものである。

【0023】経路制御ヘッダの挿入／削除の判断及び処理は、マイクロプロセッサ(MPU)1-3により実行する。中継装置のアドレスリストの挿入／削除は、ドメイン内の各中継装置にのみ適用されるローカル規定に従って実行される。明示的な中継経路を指定する中継装置アドレスリストは、ルーティングテーブル(SSRAM)1-1に格納されている。

【0024】4ポートの回線から入力されるユーザのIPv6パケットは、物理層インタフェース部1-4を経てMAC(Media Access Control)処理部1-5により終端された後、データバッファ1-6に格納される。マイクロプロセッサ(MPU)1-3は、データバッファ1-6に格納されたユーザのIPv6パケットについて、プログラムメモリ1-8に格納された処理プログラムに従ってヘッダ検査処理を実行し、必要なヘッダ情報

を読み取り、経路制御ヘッダの挿入／削除の処理を判定する。

【0025】そして、マイクロプロセッサ(MPU)1-3は、ヘッダ情報から宛先アドレスを抽出し、出側方路を決定するために該宛先アドレスを検索マシン1-2に転送する。そして、その検索結果として、経路制御ヘッダへの追加アドレスリストが格納されたメモリアドレスポインタを含む応答が返送された場合、マイクロプロセッサ(MPU)1-3は、該アドレスリストの内容をユーザのIPv6パケットに挿入する処理を行う。

【0026】なお、プログラムメモリ1-8は、データ及びコマンド用のメモリ、リクエスト及びステータス用のメモリ、及び統計情報収集用のメモリ等として使用される。また、ポリシエンジン1-7は、インテリジェントな統合パケットクラシフィケーション機能を提供し、スライサ1-9は、IPパケットとATMセルとの変換処理機能を有する。

【0027】図2は上述の経路制御ヘッダ情報をユーザのIPv6パケットに挿入する処理フローを示す。また、図3は挿入された経路制御ヘッダ情報を出側の中継装置において削除する処理フローを示す。また、図4にIPv6ヘッダフォーマットを、図5に経路制御ヘッダ情報の挿入位置を示すポインタを示す。

【0028】経路制御ヘッダ情報をIPv6パケットに挿入する処理を図2を参照して説明すると、IPv6ドメインの入口エッジ中継装置のルータは、トラフィック統計情報を受信する一方、回線から入力されたユーザのIPv6パケットの宛先アドレス(DA)によりルーティングテーブルを検索する(2-1)。

【0029】そして上記の検索結果に、LRTE(Loose Route Traffic Engineering:指定経路の中継装置アドレスを間引いて設定するトラフィックエンジニアリング)実行有りを示すデータが設定されているか否かを判定する(2-2)。なお、このデータとして、出力インタフェース番号、通過すべき中継装置アドレス等が設定されている。

【0030】上記データが設定されていた場合、受信したパケットに経路制御ヘッダが有るか否かを判定し(2-3)、経路制御ヘッダが無い場合に、新規に挿入する経路制御ヘッダをメモリにロードする(2-4)。そして、通過すべき中継装置のアドレスを検索結果からロードし、経路制御ヘッダに挿入する(2-5)。

【0031】次に、挿入した中継装置アドレスの挿入位置を示すポインタを編集し(2-6)、経路制御ヘッダと挿入位置ポインタとを受信パケットに挿入し(2-7)、IPv6ヘッダの次ヘッダフィールド値を経路制御ヘッダの次ヘッダフィールドにコピーし、その後、IPv6ヘッダの次ヘッダフィールドに経路制御ヘッダを示す値を書き込む(2-8)。そして、ルーティングテーブルの検索結果に基づいて、受信IPv6パケットを

転送（フォワーディング）する（2-9）。

【0032】次に、経路制御ヘッダ情報の削除処理を図3を参照して説明すると、IPv6/MPLSドメインの出口エッジ中継装置のルータは、経路制御ヘッダによるトラフィックエンジニアリング（TE）実行中のドメインからのIPv6パケットを受信すると（3-1）、経路制御ヘッダへの中継装置アドレス挿入位置ポインタが0か否かを判定し（3-2）、0以外である場合、中継装置アドレス挿入位置ポインタから、アドレスリスト削除か経路制御ヘッダの削除かを判定する（3-3）。

【0033】アドレスリスト削除である場合、経路制御ヘッダへの挿入位置ポインタ情報により、経路制御ヘッダに挿入されたアドレスリストを削除する（3-4）。経路制御ヘッダの削除である場合、経路制御ヘッダへの挿入位置ポインタ情報により、経路制御ヘッダに挿入された経路制御ヘッダそのものを削除し（3-5）、経路制御ヘッダの前段のヘッダに有る次ヘッダフィールドに、削除した経路制御ヘッダに付与されていた次ヘッダフィールド値を書き込む（3-6）。そして、経路制御ヘッダ情報を削除した分、ペイロード長を減算してIPv6ヘッダに書き込み（3-7）、次経路へのデータ転送（フォワーディング）を開始する（3-8）。

【0034】IPv6のヘッダフォーマットは図4に示すように、バージョン、トラフィッククラス（優先度）、フローラベル、ペイロード長、次ヘッダ、ホップ制限、送信元アドレス及び宛先アドレスを含む基本ヘッダと、ホップバイホップオプションヘッダ、宛先オプションヘッダ、ルーティングヘッダ、フラグメントヘッダ、認証ヘッダ、及び上位レイヤヘッダ等を含む拡張ヘッダとから成る。

【0035】上記拡張ヘッダのルーティングヘッダに通過経路の中継局アドレスが書き込まれる。該ルーティングヘッダのフォーマットを図5に示す。同図はタイプ0のルーティングヘッダのフォーマットを示し、ルーティングヘッダには、次ヘッダ、拡張ヘッダ長、ルーティングタイプ、残セグメント、予備（Reserved）フィールド、通過経路の各中継装置アドレスの領域が備えられている。

【0036】次ヘッダのフィールドは8ビットのセクタであり、ルーティングヘッダの直後に続くヘッダのタイプを示し、IPv4パケットにおけるプロトコルフィールドと同様の値が使用される。拡張ヘッダ長のフィールドは、8ビットの符号無し整数によりルーティングヘッダの8オクテット単位の長さを表す（但し、先頭の8オクテット分は含まない）。なお、タイプ0のルーティングヘッダの拡張ヘッダ長の値は、ルーティングヘッダ内の中継装置アドレス数の2倍の値に等しくなる。

【0037】ルーティングタイプのフィールドには8ビットでルーティングタイプ値（この場合“0”）が書込まれる。残セグメントのフィールドには、8ビットの符

号無し整数で、経路上の残りのセグメント数、即ち、最終宛先に到達するまでに経由するように明示されている中間ノード（中継装置）の数が書込まれる。

【0038】予備（Reserved）フィールドは32ビットの情報フィールドで、送信時に“0”に初期化され、受信側では無視される。通過経路の各中継装置アドレス[1]…[n]は、1からnまで番号付けされた各中継装置の128ビット長のアドレスである。

【0039】本発明の第1の実施形態は、経路制御ヘッダへの中継装置アドレスの挿入位置を示すポインタとして、予備（Reserved）フィールドに、一例として以下のような制御フィールドのフォーマットを挿入する。

（1）該予備（Reserved）フィールドに、マルチプロトコラベルスイッチング（MPLS）方式におけるシム（shim）ヘッダ形式のフィールドを定義し挿入する。

（2）シム（shim）ヘッダの段数は1段とする。

（3）予備（Reserved）フィールドのデータが“0”でなければ、シム（shim）ヘッダが有意データとして挿入されていると判断する。

（4）この場合、シム（shim）ヘッダのラベルフィールド（20ビット）を2オクテットと4ビットのサブフィールドに分割し、経路制御ヘッダによるトラフィックエンジニアリング時の挿入経路リストの範囲指定を記述する。

【0040】図6にシム（shim）ヘッダフォーマットを示す。該シム（shim）ヘッダフォーマットは、RFC3032 Jan. 2001により規定されたものである。シム（shim）ヘッダフォーマットは、20ビットのラベル（Label）フィールド、3ビットの実験使用（Exp）フィールド、1ビットのボトムスタック（S）フィールド、8ビットの生存時間（TTL）フィールドの32ビットから構成される。

【0041】ラベル（Label）フィールド内の第1の8ビットサブフィールド（sub-label-1）には、トラフィックエンジニアリング用の経路制御ヘッダ内容を挿入する位置を示すポインタを記述する。また、第2の8ビットサブフィールド（sub-label-2）には、トラフィックエンジニアリング用の経路制御ヘッダ内容を挿入する数を示すポインタを記述する。ラベル（Label）フィールド内の4ビットサブフィールド（sub-label-3）は予備（Reserved）である。

【0042】このように、パス設定制御プロトコルによってパス設定処理を行わずに、IPv6パケットの経路制御ヘッダに経路変更の制御機構を組み込むことにより、別の経路へユーザのIPv6パケットを迂回させる。この場合、明示的経路指定情報を他の中継装置群に通知するために、本発明の第1の実施形態は、拡張ヘッダフィールドに含まれるTLV（Type/Length/Value）群のフィールドである経路制御ヘッダの予備（Reserved）フィールド中に、新たに追加した経路（中継装置ア

ドレスリスト)の挿入位置ポインタと中継指定段数とを記述する。

【0043】この場合、該経路変更の制御機構をサポートしていない中継装置において、経路制御ヘッダの予備(Reserved)フィールドを“0”にクリアする処理を実行するものがあつたりすると、新たに追加した経路の削除処理を出口エッジの中継装置で行うことができなくなり、動作に不具合が発生する。

【0044】即ち、中継装置アドレスリストが付加挿入されている旨の表示を、経路制御ヘッダ中の予備(Reserved)フィールドにより行って経路変更の制御を行うこととしたが、本来、予備(Reserved)フィールドは、将来の使用のために確保されているフィールドであり、他のIPv6パケット中継機構の実装において、この予備(Reserved)フィールドをゼロクリアして中継したり、その内容をチェックしてゼロ以外ならば、異常検出としてアラーム処理したり、といった様々の判断及び処理等に使用されることが予想される。

【0045】そこで、経路上の中継装置が常に検査すべきヘッダとして規定され、他の用途にも使用可能な汎用ヘッダフィールドであるホップバイホップオプションフィールドに新規にフィールド規定を追加し、中継装置アドレスリストの挿入を表示することができる。ホップバイホップオプションフィールドとして規定されたものは、フィールド値の一部であつて、それ以外は未規定である。

【0046】この規定済み部分の値として、オプションタイプの上位から3ビット目は、このオプションデータがパケットの最終宛先までの途中で修正することができるかどうかを指定するビットであり、このビットを、オプションデータを途中で変更できない値(0)に設定することにより、前述のゼロクリアによる不具合を解決することができる。但し、ホップバイホップオプションヘッダの挿入削除処理の負荷は増大する。

【0047】しかしながら、予備(Reserved)フィールド使用による不具合を回避するために、経路制御ヘッダの予備(Reserved)フィールドの代わりに、本発明の第2の実施形態として、IPv6プロトコル標準規定のホップバイホップオプションヘッダのパッドNオプション機能を拡張し、新たに追加した経路の挿入位置表示ポインタ及び中継段数指定を表示せしめることとする。

【0048】IPv6プロトコル標準仕様におけるパッドNオプション(alignment requirement: none)のフィールド定義を以下に説明する。パッドNオプションは、ヘッダのオプションエリアに2オクテット以上のパディングを挿入するのに使用される。図7の(a)にパッドNオプション(alignment requirement: none)フィールドのフォーマットを示す。Nオクテットのパディングの場合、オプションデータ長フィールドに(N-2)の値を記述し、オプションデータフィールドには

(N-2)オクテットの“0”値が記述される。

【0049】IPv6プロトコル標準仕様の規定により、ヘッダ内の一連のオプションはヘッダ内に現れた順に厳密に処理しなければならない。例えば、ヘッダ内の特定のオプションを探して、それに先行する全てのオプションを処理する前に、そのオプションの処理してはならないとされている。なお、パッド1オプションは、ヘッダのオプションエリアに1オクテットのパディングを挿入するのに使用され、1オクテットを超えるパディングが必要な場合には、複数のパッド1オプションを使用するよりも、ここに示したパッドNオプションが使用される。

【0050】オプションタイプ識別子は、単にオプションの特定のタイプを識別する以上の役割を有し、その上位2ビットは、処理中のIPv6ノードがそのオプションタイプを認識することができない場合にどのように扱うかを指定するように符号化されている。オプションタイプ上位2ビットの符号の意味は図7の(b)に示すように規定され、上記パッド1オプションとパッドNオプションは図7(b)の表の符号“00”の規定に従う。中継装置がホップバイホップオプションヘッダ付のIPv6パケットを受信すると、そのオプションタイプを検査し、上記パッド1オプションかパッドNオプションであれば、それに対応した処理を行ってから次の拡張ヘッダ処理を行なう。

【0051】ここで本発明における追加経路(中継装置アドレスリスト)の挿入位置ポインタと中継指定段数の記述に関して、上記パッドNオプションのオプションデータ部分に一例としてマルチプロトコルラベルスイッチング(MPLS)方式のシム(shim)ヘッダを挿入する。シム(shim)ヘッダの段数は、オプションデータ長フィールドを参照することにより、4オクテットの何倍かを計算すればシム(shim)段数を判別することができる。

【0052】オプションタイプ値は、“00000XX”を使用し、オプションデータが“0”でなければ、有意データとしてシム(shim)ヘッダが挿入されていると判断する。この場合、シム(shim)ヘッダのラベルフィールド(20ビット)を2オクテットと4ビットのサブフィールドに分割し、経路制御ヘッダによるトラフィックエンジニアリング時の挿入経路リストの範囲指定を記述する。

【0053】ホップバイホップオプションヘッダ付のIPv6パケットを受信した場合、これに付加する形で経路制御ヘッダ情報の挿入位置ポインタを記述する。ホップバイホップオプションヘッダの無いユーザIPv6パケットを受信した場合、新規にホップバイホップオプションヘッダを設定する。

【0054】この場合、挿入する経路制御ヘッダの前の次ヘッダフィールドには、ホップバイホップオプション

10

20

30

40

50

ヘッダである旨の表示に書き換える必要がある。このパッドNオプション内に、図6に示したような経路制御ヘッダの挿入位置情報ポインタを記述する。この位置情報ポインタのフォーマットは、一例として図6に示したシム(shim)ヘッダフォーマットを流用することができる。

【0055】次に、IPv6中継システムのドメイン内において、次ホップの中継装置(ルータ)が運用停止している場合の処理フロー例を図8に示す。各中継装置は、通常、OSPF(Open Shortest Path First)に代

表される動的なルーティングプロトコルを動作させているので、隣接した中継装置が運用中か否かを検出することができる。

【0056】そこで、経路制御ヘッダによるトラフィックエンジニアリング実施ドメインからのIPv6パケットを受信すると(8-1)、IPv6パケットのヘッダに記述されたアドレスの中継装置が正常に運転されているかを判定し(8-2)、障害等により運転を停止している場合、その中継装置への経路情報が経路制御ヘッダ(ルーティングテーブル)に記述されていたとしても、経路制御ヘッダに記述されている次の中継アドレスとIPv6基本ヘッダの宛先アドレスとを交換する処理を実行し(8-3)、次経路へのIPv6パケット転送を実行し(8-4)、ユーザのIPv6パケットの中継を継続する。

【0057】このように、エンドホスト装置が知り得ない中継処理の途中段階において、中継装置が自律的に経路制御ヘッダの値を設定して、能動的にネットワーク内の輻輳緩和のためにユーザIPv6パケットに付与された元々の経路制御ヘッダに、又は全く新規に経路制御ヘッダの挿入削除処理を行う。この場合、そのネットワークドメイン内の中継装置で、障害を起こして中継動作を停止しているものの直前ホップの中継装置が、停止中の中継装置を認識すると(Keep aliveパケットの送受で判別可能)、その中継装置において経路制御ヘッダのアドレスリスト入れ替え処理を実施し、次の宛先アドレス(運用中の経路)に向けてユーザIPv6パケットを中継する。

【0058】ユーザが設定した経路制御ヘッダに関しては、通常このような処理は行われない。宛先到達不可のエラーメッセージが送信元装置に返送され、送信元装置の処理により再送されることとなる。しかし、本発明において、経路途中の中継装置で経路制御ヘッダを作成して挿入するので、再送制御を送信元ホスト装置に委ねることなく、経路制御ヘッダの内容を運用中の宛先アドレスに置換する処理を自律的に実行することで、送信元ホスト装置に処理負荷を掛けることなく、ユーザIPv6パケットの中継を継続することができる。

【0059】次に、入力インタフェース及び出力インタフェースのメッセージ転送容量を基に経路制御情報の挿

入制御を行う処理について図9を参照して説明する。本発明による明示的経路指定中継装置は、中継経路のメッセージ転送容量(MTU)サイズの違いを意識しなければならない。

【0060】このために、中継装置の保有するルーティングテーブルに通常記載のある“Interface”フィールドの内容を検査することにより、中継経路のメッセージ転送容量(MTU)サイズを基に、経路制御情報の付加挿入処理の制御を行う。具体的には、“Interface”フィールドに、例えば、イーサネット(登録商標)、FDDI(Fiber Distributed Data Interface)、ATM(Asynchronous Transfer Mode)等のインタフェース種別(リンク種別)が記述されているので、これを参照してメッセージ転送容量(MTU)サイズの大小関係を判別することができる。

【0061】IPv6/マルチプロトコルラベルスイッチング(MPLS)システムにおけるドメイン入口エッジのラベルスイッチングルータ(LSR)は、入側インタフェースのメッセージ転送容量(MTU)のサイズと、出側インタフェースのメッセージ転送容量(MTU)のサイズを、ルーティングテーブルのインタフェース種別(リンク種別)情報を参照して比較し大小関係を判断する。

【0062】図9に示すように、まず、ユーザIPv6パケットの受信側インタフェースのメッセージ転送容量(MTU)を記憶する(9-1)。次に、ルーティングテーブルを検索し、IPv6経路制御ヘッダでの負荷分散処理を行う出側インタフェースの種別を判定する(9-2)。そして、受信側インタフェースのメッセージ転送容量(MTU)より、出側インタフェースのメッセージ転送容量(MTU)のサイズが大きいかどうかを判定する(9-3)。

【0063】受信側インタフェースより出側インタフェースのメッセージ転送容量(MTU)のサイズが小さいか等しい場合、当該インタフェースでは経路制御ヘッダでの分散を禁止するフラグを設定する(9-4)。中継装置はこの禁止フラグが設定されている場合、ルーティングテーブルの検索結果に経路制御ヘッダへの追加アドレスリストを設定する処理を中止する。その結果、当該インタフェース上では経路制御ヘッダによる経路指定制御は実行されない。

【0064】ルーティングテーブルの保持情報の例を図10に示す。ここで、“Interface”と記述されたフィールドはインタフェース種別(リンク種別)を示し、図10において“ei31”の記述から、インタフェース種別(リンク種別)がイーサネット(登録商標)であることが判別される。

【0065】そこで、ユーザのIPv6パケットを受信したインタフェース種別と、ルーティングテーブルを検索して得られたインタフェース種別(Interface

eフィールドの内容)とを比較して、出口インタフェースのメッセージ転送容量(MTU)の方が大きくなければ、経路制御ヘッダへのアドレス追加処理を行わないようにする。

【0066】(付記1) 複数の中継装置を経由してデータを中継するネットワークを構成する中継装置であって、迂回中継の経路を指定するための明示的経路設定制御手段を具備する中継装置において、自中継装置の収容する回線インタフェースのトラフィック状況を監視し、統計情報として収集する手段と、周囲の他の中継装置に該統計情報を通知する手段と、他の中継装置から通知された該統計情報及び自中継装置に保有しているネットワーク経路情報を基に、迂回中継するか否かを判断する手段と、迂回中継の経路を指定するための明示的経路設定制御手段とを備え、前記明示的経路設定制御手段は、当該中継装置がネットワークドメインの入口エッジに位置する場合、明示的経路指定用の経路制御ヘッダに、該ドメイン内でのみ使用する明示的経路指定情報としての中継装置アドレスリストを付加挿入し、付加した経路の挿入位置ポインタ及び中継指定段数を、インターネットプロトコルの拡張ヘッダ内のフィールドに記述する手段と、当該中継装置がネットワークドメインの出口エッジに位置する場合、入口エッジ中継装置において付加挿入された明示的経路指定情報としての中継装置アドレスリスト情報を削除する手段と、を具備することを特徴とする明示的経路指定中継装置。(1)

(付記2) 前記ネットワークドメイン内の運用停止中の中継装置を認識する手段を備え、前記明示的経路指定情報として記述された中継装置アドレスリスト中に、該運用停止中の中継装置が存在する場合、前記経路制御ヘッダの中継装置アドレスを、運用中の中継装置への経路に入れ替える手段を備えたことを特徴とする付記1に記載の明示的経路指定中継装置。(2)

(付記3) 入力したパケットデータの中継経路上のメッセージ転送容量が、入力インタフェース上のメッセージ転送容量より大であるか否かを事前に検出する手段を具備し、出力インタフェースのメッセージ転送容量が入力インタフェースのメッセージ転送容量より小さい場合、前記経路制御ヘッダの中継装置アドレスリストの付加挿入を行わないように制御する手段を備えたことを特徴とする付記1又は2に記載の明示的経路指定中継装置。(3)

(付記4) 前記明示的経路設定制御手段は、付加した経路の挿入位置ポインタ及び中継指定段数を、インターネットプロトコルの拡張ヘッダ内の経路制御ヘッダ内の予備フィールドに記述する手段を備えたことを特徴とする付記1に記載の明示的経路指定中継装置。

(付記5) 前記明示的経路設定制御手段は、付加した経路の挿入位置ポインタ及び中継指定段数を、インターネットプロトコルの拡張ヘッダ内のホップバイホップオ

プションヘッダフィールドに記述する手段を備えたことを特徴とする付記1に記載の明示的経路指定中継装置。

【0067】

【発明の効果】以上説明したように、本発明によれば、トラフィックの集中を回避する負荷分散制御を行う際に、変更する経路の中継装置アドレスリストを経路制御ヘッダに対して挿入／削除することにより、中継装置内のルーティングテーブルの形式等を一切変更することなく、また、中継装置間でのパス設定制御用のLDP/CLR-LDP又はRSVP等のプロトコルを用いることによるオーバーヘッドや処理の遅延を発生させることなく、また他にシグナリング機能等を用いることなく、トラフィック負荷をネットワーク内に分散中継するための経路変更制御することが可能となる。

【0068】本発明によるトラフィックエンジニアリング実行中のドメイン内で、経路制御ヘッダに一時的に中継装置アドレスが追加されたとしても、ドメインの出口エッジで追加中継装置アドレスを削除することにより、他の中継装置では一切問題を生ずることなく、トラフィック負荷を分散中継することができる。

【0069】また、IPv6パケットを中継する装置側では、通常どおりのヘッダチェックを行い、該経路制御ヘッダに記述された経路に従ってIPv6ユーザパケットを中継させることが可能である。この結果、ネットワーク内に輻輳が発生していたリンクへの入力負荷を速やかに減少させることができ、この経路を利用していたインターネット利用中のエンドユーザの通信スループットを見かけ上増大することができ、タイムラグの無い速やかな経路切替えを、インターネットドラフトやRFCに記載されたIPv6プロトコル処理に付加する形で、標準のプロトコル仕様との矛盾を発生することなく、実現することができる。

【0070】また、運用停止中の中継装置が経路制御ヘッダの指定経路上に存在する場合、該運用停止中の中継装置への経路を、運用中の中継装置への経路に入れ替えることにより、宛先ホスト装置へ到達不可のエラーメッセージが返送されたり、又は何のエラーメッセージも返送されずにIPv6パケットが廃棄されることによる中継処理の中断動作を生じることなく、通信を継続させることができる。

【図面の簡単な説明】

【図1】本発明による明示的経路指定中継装置の機能ブロック図である。

【図2】経路制御ヘッダ情報を挿入する処理のフローを示す図である。

【図3】挿入された経路制御ヘッダ情報を削除する処理のフローを示す図である。

【図4】IPv6のヘッダフォーマットを示す図である。

【図5】ルーティングヘッダのフォーマットを示す図で

ある。

【図6】シム (shim) ヘッダフォーマットを示す図である。

【図7】パッドNオプションフィールドのフォーマット及びオプションタイプの符号内容を示す図である。

【図8】次ホップの中継装置が運用停止している場合の処理のフローを示す図である。

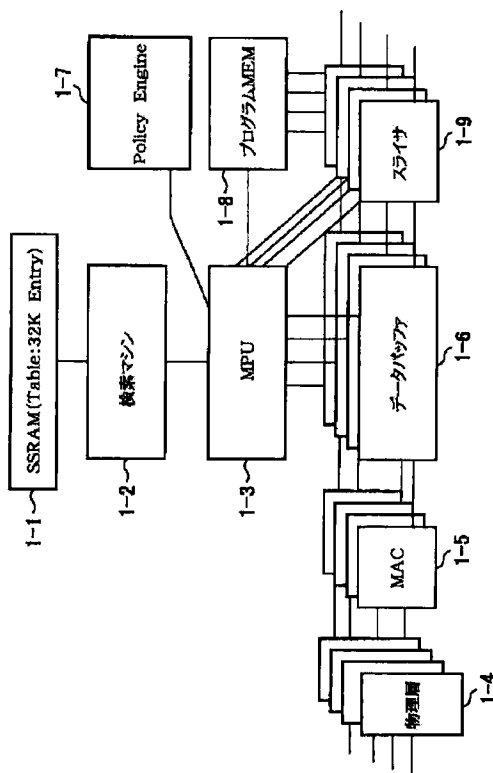
【図9】入力インタフェース及び出力インタフェースのメッセージ転送容量を基に経路制御情報の挿入制御を行う処理のフロー図である。

【図10】ルーティングテーブルの保持情報の例を示す図である。

【図11】従来の明示的ルーティングを行う中継装置の機能構成を示す図である。

【図1】

本発明による明示的経路指定中継装置の機能ブロック



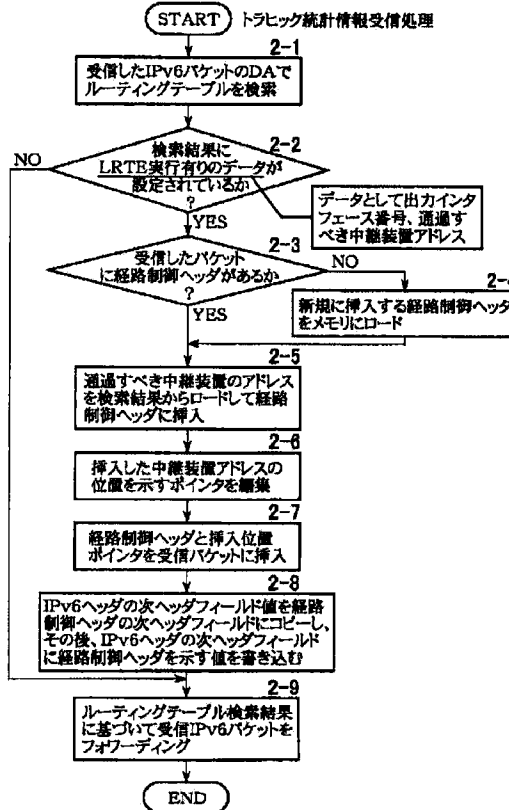
【図12】マルチプロトコルラベルスイッチング (MPLS) ネットワークにおける負荷分散中継の動作例を示す図である。

【符号の説明】

- 1-1 ルーティングテーブル (SSRAM)
- 1-2 検索マシン
- 1-3 マイクロプロセッサ (MPU)
- 1-4 物理層インタフェース部
- 1-5 MAC (Media Access Control) 処理部
- 1-6 データバッファ
- 1-7 ポリシエンジン
- 1-8 プログラムメモリ
- 1-9 スライサ

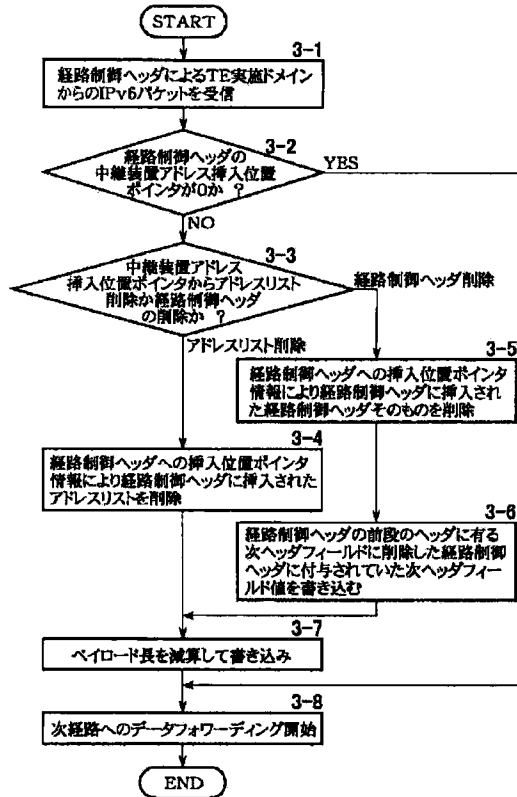
【図2】

経路制御ヘッダ情報を挿入する処理のフロー



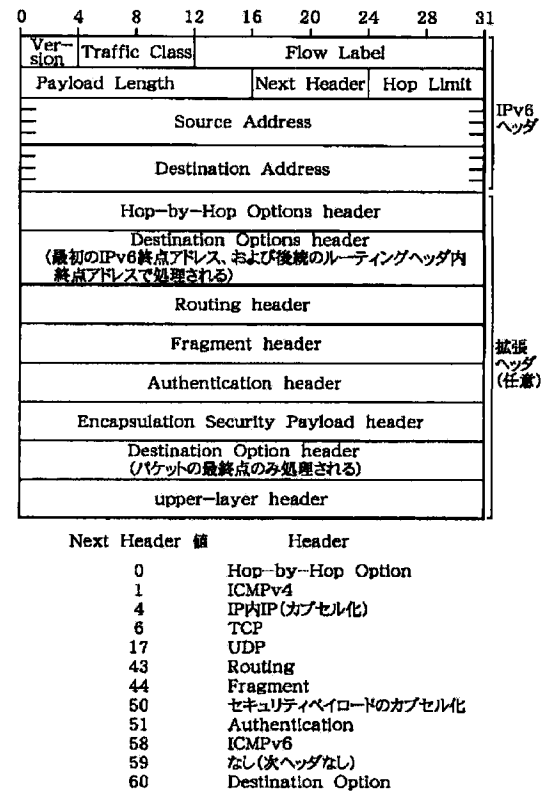
【図3】

挿入された経路制御ヘッダ情報を削除する処理のフロー



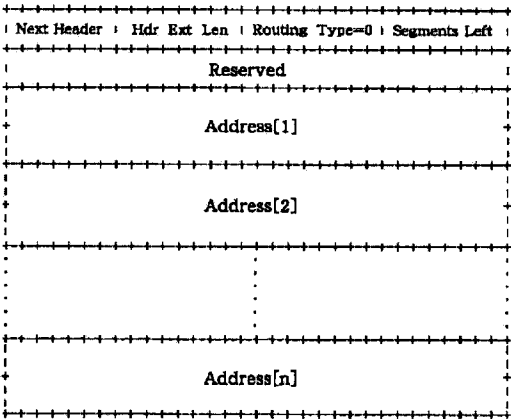
【図4】

IPv6のヘッダフォーマット



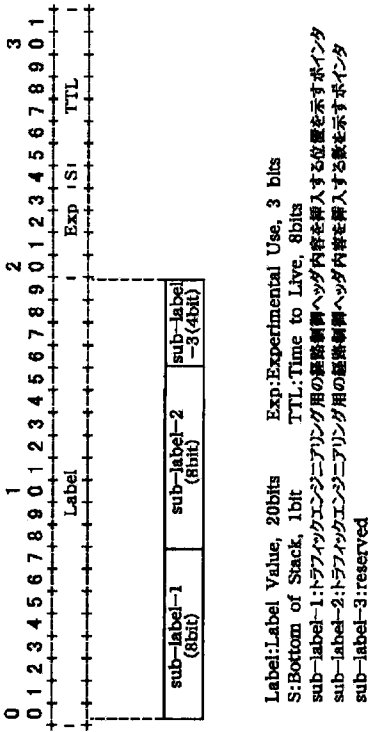
【図5】

ルーティングヘッダのフォーマット



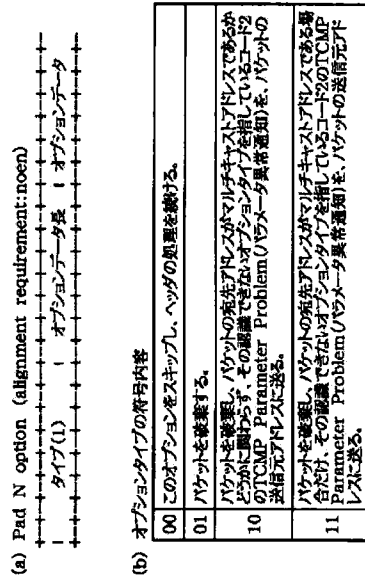
【図6】

シム (shim) ヘッダフォーマット



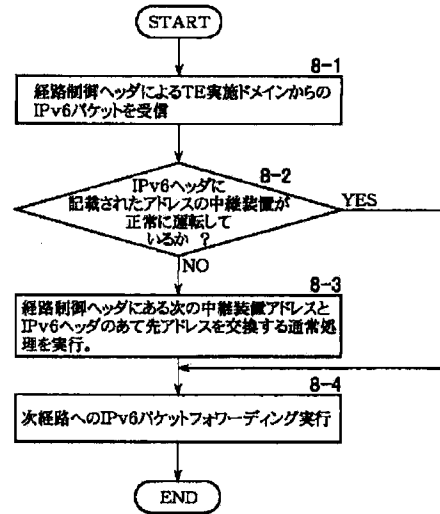
【図7】

パッドNオプションフィールドのフォーマット及びオプション
タイプの符号内容



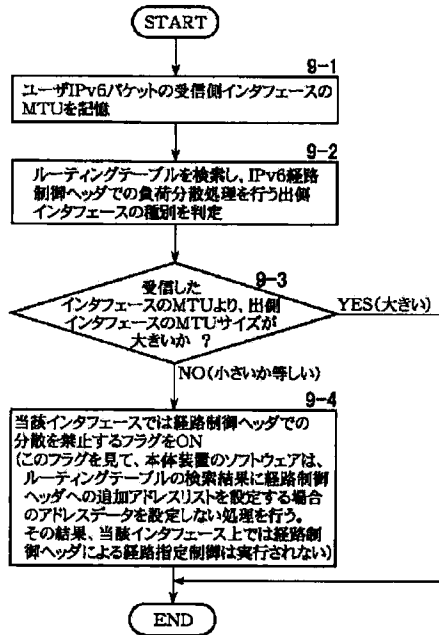
【図8】

次ホップの中継装置が運用停止している場合の処理のフロー



【図9】

メッセージ転送容量を基に経路制御情報の挿入制御を行う処理のフロー



【図10】

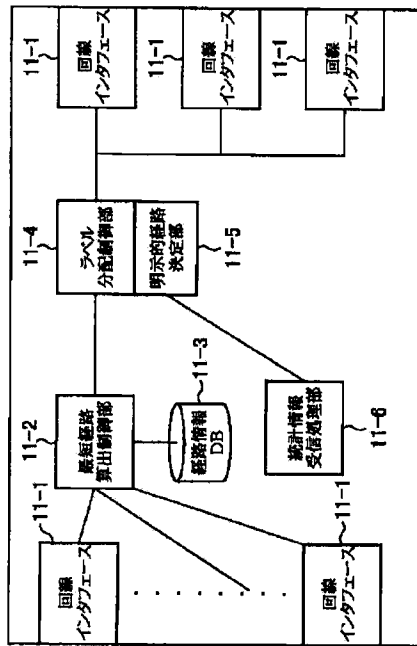
ルーティングテーブルの保持情報の例

Routing table for Internet on slot0

Destination	Gateway	Flags	Refcnt	Use	Interface
0.0.0.0	110.257.128.21	UG	0	0	ei31
110.146.1.1	110.257.128.21	UGH	0	0	ei31
110.146.1.2	110.257.128.21	UGH	0	0	ei31
110.146.1.3	110.257.128.21	UGH	0	0	ei31
110.146.1.4	110.257.128.21	UGH	0	0	ei31
110.243.88.0	110.257.128.21	UG	0	0	ei31
110.243.90.0	110.257.128.21	UG	0	0	ei31
110.261.205.2	110.257.128.21	UGH	0	0	ei31
110.261.206.1	110.257.128.21	UGH	0	0	ei31
127.0.0.1	127.0.0.1	UH	0	0	lo
224.0.0.9	127.0.0.1	UH	0	2	lo

【図11】

従来の明示的ルーティングを行う中継装置の機能構成



【図12】

マルチプロトコラベルスイッチング (MPLS) ネットワークにおける負荷分散中継の動作例

